Table of Contents

Hepatitis 'C' Status File

This file contains one record per subject and their hepatitis C status. This is the SUBJECT'S overall status, based on laboratory information.

NTT1SWY.CS5.HCVSTAT.SASDS(.STAT)

*** One record per subject, sorted by SUBJ_ID.

Input Files

UAT1PIV.HIV9.HEP.SASDS (.SCREEN) & (.BLOT)

Variables

SUBJ_ID	-	Subject Id Number
HCV_STAT	-	Overall status
TEST	-	Type of Hepatitis C test used

HCV_STAT is set as follows using PCR/RIBA and screen results:
Only the '+' PCR results are considered

PCR	RIBA2	RIBA3	HCV Status
+	any	any	+
	+	-	+
+		/	+
	+	+	+
	-	-	-
	-	/	/
	-	+	+
	/	-	-
	/	/	/
	/	+	+

If no PCR or RIBA was done, the subject takes the status from the screen result. If the screens conflict the status is '/'.

Updating the File

Run from NTT1SWY.CS5.HCV.SAS(CRESTAT)

HCV Genotype/Simmonds' Analysis File

File NTT1SWY.CS5.HCVGENOS.SSD(.HCV)
C:\hemfiles\allgeno.sd2

Duplicate records issue

The file was originally designed to have one record per subject
Dr. Eyester's original data was added and contained paired samples - which created duplicates for subjects sometimes with conflicting genotype results
.
These duplicate samples must be handled by the specs of individual analysis.

About Simmonds' HGV Data

The RFLP data from Dr. Simmonds also contained an HGV flag. Since the only HGV data we maintain is the HGV Couples File which will be analyzed as a unit, Dr. Simmond's HGV data will temporatily be kept here.

About the genotypes

This file contains the HCV genotype results from Dr. Hatzakis at the Hemophilia Center in Athens. The genotypes come in as text and reflect combinations of genotypes 1-5, Positive NOS, or Negative. It also contains 110 HCV RFLP results from Dr. Simmonds in Scotland. 42 of these were sent directly from Dr. Eyster. These 42 records will not have genotype information unless by chance.

An few examples are these;

1a	-	A clean genotyping of 1a
1a, 1b	-	A mixture of genotype 1; that is 1a AND genotype 1b detected
1, 1a	-	A different mixture: that is '1' (not genotyped further) AND 1a
2a/2c	-	EITHER 2a OR 2c detected
1a, 2a/2c, 4		Genotype 1a, EITHER 2a or 2c, AND Genotype 4

.

About the recoding

They are recoded 2 ways for analysis

The 'ANY' flags ANY1-ANY5 (for genotypes 1-5) ANY0 (negative) and ANY9 (positive NOS) are set independently of each other. For example, genotype '1a, 4c/4d' will turn on ANY1, and ANY4..

The grouping variable '**GROUP**', however, is set as a **hierarchy** in the following order.

1 = More than 1 genotype (ie 1a, 4c)
2 = Negative
3 = Positive - not otherwise specified
4 = Any 1 (for example genotype '1' or genotype '1, 1a')
5 = Any 2 '' for genotype 2
6 = Any 3 '' for genotype 3
7 = Any 4 '' for genotype 4
8 = Any 5 '' for genotype 5

When Jim talks about 'ANY1' for example, he is referring to the hierarchy of variable GROUP. The ANY flags are just internal to the file for coding and checking purposes..

Variables

ANY0	-	Genotype result is 'NEGATIVE'
ANY1	-	At least 1 Genotype 1
ANY2	-	At least 1 Genotype 2
ANY3	-	At least 1 Genotype 3
ANY4	-	At least 1 Genotype 4
ANY5	-	At least 1 Genotype 5
ANY9	-	Genotype result is POSITIVE but UNSPECIFIED
BIRTH	-	Birth date (SAS) 1 = 1900-1949 2 = 1950-1954 3 = 1955-1959 4 = 1960-1964 5 = 1965-1969 6 = 1970-1974 7 = 1975-1979 8 = 1980-1984 9 = 1985-1996
COUNTRY	-	US or European country
DRAW	-	Draw date of genotyped sample - SAS
DT_STAMP	-	Date the GENOTYPE data was received - (SAS) records with RFLP results only have a missing date stamp
GENOTYPE	-	Genotype
GROUP	-	Genotype group 1 = More than 1 genotype 2 = Negative 3 = Positive - not otherwise specified 4 = Any 1 5 = Any 2 6 = Any 3 7 = Any 4

		8 = Any 5
GROUPTXT	-	Probably silly to include, but comes in handy in printouts More than 1 Neg '+' NOS ANY 1 ANY 2 ANY 3 ANY 4 ANY 5
HCV_STAT	-	HCV Status **at time of selection for genotyping**
HGVMAT	-	Material code of Simmond's HGV results
LIVERDT	-	Liver Disease Dx date - SAS
LIVFLAG	-	Liver Disease flag 0/1/2 (neg/pos/possible)
RFLP	-	RFLP result from Dr. Simmonds
SAMPL_ID	-	Sample Id
SIMMHGV	-	Simmond's HGV result
STATANL	-	HIV Status **at time of selection for genotyping**
SUBJ_ID	-	Subject Id

Comments about the variables on the file

The US subjects selected to be genotyped were chosen by **HCV status** (all positive), their **HIV status** (approximately 1/2 and 1/2) and **birthyear** (cross section) at the time of selection. The Greeks that Dr. Hatzakis selected were all HIV positive, all HCV positive (by his own test) and not selected by birth year. 5 Greek subjects were HCV neg/equiv by our test. In order to match analysis already performed, the HIV, HCV status flags and birth year were kept on the file. The danger, of course is that these could change by data correction or more tests, so the changable nature of these variables must be kept in mind.

All of the Greek subjects are in the MHCS study. However the results on 25 Greek subjects were on **Subject Id/Draw Date combinations** that are not in our background file. Dr. Hatzakis only provided the subject Id and date, therefore the Sample Id on these records is missing. A report to this effect was sent to him and he verified the dates. This has already resulted in analysis problems when the draw dates were later than our LKA. Dr. Hatzakis has given us updated LKA's and AIDS status. This will be incorporated into our permanent files.

Group and Genotype Combinatins

GROUP	GROUPTXT	GENOTYPE
1	More than 1	1, 1a, 2a/2c, 2b

			1,2,3
			1a, 2a/2c
			1a, 2a/2c, 4
			1a, 2b
			1a, 4c/4d
			1b, 2b
			1b, 3a
			3a, 2b
			4c/4d, 1a
2	Neg		Neg
3	'+' NOS		Pos NOS
4	ANY 1	1	
			1, 1a
			1, 1a, 1b
			1a
			1a, 1b
			1b
			1c/1d
5	ANY 2	2a	
			2a,2b
			2a/2c
			2a/2c, 2b
			2b
6	ANY 3	3a	
			3a,3b
7	ANY 4		4a
			4c/4d
8	ANY 5	5a	

Input Files and DATE STAMP issues

1. NTT1SWY.CS5.HCVSUB.GENO0197.SSD
Hershey only DATE STAMP - 1/1/97 n=58
The create program is &NTT1SWY.HCV.GENO98.SAS(MAKE0197) and c:\hcv\hcvsas\make0197 see the following section entitled **First Hepatitis Sub-Genotype File** for details of this data.l

2. NTT1SWY.CS5.HCVSUB.GENO0797.SSD
All U.S. subjects.
Large group from Dr. Hatzakis is the core subjects of the 'HCV Combined file
DATE STAMP - 7/1/97 n=320
The create program is lost - but information about the input files is in program &NTT1SWY.HCV.GENO97.SAS(MAKEFILE)

3. NTT1SWY.CS5.HCVSUB.GENO0198.SSD
Additonal results from Dr. Hatzakis. Greeks only.
DATE STAMP - 1/1/98 n=129
The create program is &NTT1SWY.HCV.GENO98.SAS(MAKE0198) and c:\hcv\hcvsas\make0198

4. NTT1SWY.HCV.SIMMONDS.RFLP.L26
The data was created from diskette of results
DATE STAMP - MISSING on results sent direction from Dr. Eyster n=42

5. NTT1SWY.CS5.HCVSUB.GENO0998.SSD
All Hershey subjects
New and old data from Hershey. Replaces the previously received data.
DATE STAMP - 9/11/98 on completely new records. File contains new data, duplicates of old data, and updated data on previously existing subject/draws.
Create program is saved as c:\hcv\hcvsas\make0998.sas and &NTT1SWY.HCV.GENO98.SAS(MAKE0998)

6. NTT1SWY.CS5.HCVSUB.GENO1298.SSD
New data from Dr. Hatzakis
DATE STAMP - 12/01/98
Create program is saved as c:\hcv\hcvsas\make1298.sas and &NTT1SWY.HCV.GENO98.SAS(MAKE1298)

Updating the File Run from NTT1SWY.HCV.GENO98.SAS(CREANAL).

HCV Birth Cohort File (Combined File)

(Combined Genotype/Serotype/bDNA File
File NTT1SWY.HCV.COMBINED.SSD(.HCV)
C:\hemfiles\hcvcomb.sd2

One record per sample, sorted by sample

About the file

310 of the samples genotyped at one time by Dr. Hatzakis (date stamped 070197 on the genotype file) were selected to be tested by all three HCV PCR tests; Bdna, genotype and serotype (Chiron test at SAIC). When necessary, due to shortage of blood, the samples closest to the genotyped sample was chosen and tested. The 'HAS' flags refer to the presence of test results - not whether or not the results are on the same sample. The time variables can be used to select results based on some specified minimum time from genotype draw dates.

Input Files NTT1SWY.CS5.HCVGENOS.SSD(.HCV)
UAT1PIV.HIV9.PCR.SASDS(.PCR)

Variables: *** see notes in HCV Genotype/Simmonds Analysis File**

AGEDRAW - Age at draw
AIDS - AIDS flag
ANY0 * - Genotype Neg
ANY1 * - Any 1
ANY2 * - Any 2
ANY3 * - Any 3
ANY4 * - Any 4
ANY5 * - Any 5
ANY9 * - Pos/NOS
BDNA - HCV bDNA viral load
BDNACAT - HCV bDNA category (0-3 neg-high)
BIRTH - Birth date - SAS
BRTHYR - Birth year category
CNTL1 - Control Band 1
CNTL2 - Control Band 2
CORE1 - Core1 Band4
CORE2 - Core2 Band5
COUNTRY - Country
DAIDS - AIDS date - SAS
DILUTION - Dilution
DRAW - Draw Date - SAS
DT_DEATH - Death date mmddyy
GENOTYPE - Genotype
GROUP - Genotype Group #
GROUPTXT - Genotype Group Name
LFLAG - Liver Disease flag 0/1/2 (neg/pos/possible)
LIVERDT - Liver Disease Dx date - SAS
LIVSURV - Liver Survival - Startdt --> LIVERDT
LKA - Last Known Alive - SAS
NS41 - NS4 Band1
NS42 - NS4 Band2

NS43	-	NS4 Band3
RFLP	-	RFLP - Dr. Simmonds,Scotland
SERODT	-	Seroconversion date - SAS
SEROTYP	-	Chiron Serotype
STARTDT	-	Later of 5/29/82 or birth - SAS
STATANL	-	HIV Status
SUBJ_ID	-	Subject Id
SURVIVAL	-	STARTDT - LKA
VITLSTAT	-	Vital Status 1/2 (alive/dead)

Updating the file

Run c:\hcv\hcvsas\combine.sas - it creates and uploads the file
Program also saved as &NTT1SWY.HCV.GENO98.SAS(COMBINE)

****** One time update**

The file has been modified one time since its creation. In January 1999, 3 subjects had HCV serotypings successfully refined.

The program used is saved as c:\hcv\hcvsas\combupdt.sas and
&NTT1SWY.HCV.GENO98.SAS(COMBUPDT)

Analysis File for 01/99 Publication

File NTT1SWY.HCV.PAPER99.SSD(.HCV)
C:\HCV\PAPER99\PAPER99.SD2

One record per subject, sorted by subject

Create program
&NTT1SWY.HCV.PAPER99.SAS(CREANAL)
C:\HCV\PAPER99\CREANAL.SAS

About the file

This file is the analysis file that will be frozen and used for 01/99 analysis done for publication. It contains 1192 HIV positive and 624 HIV negative subjects for a total of 1816 subjects. All are HCV positive .It uses the updated DEC98 liver disease file, and only second generation Bdna results (RSLT_TYP 'C2'). The subjects are divided into 3 groups:

*** 1 Birth Cohort - the 310 subjects in the birth cohort (combined) file
2 Plus Cases - 38 liver disease cases that, like the birth cohort file have bdna , serotyping and genotyping on the same sample
3 Every one else

Input files

NTT1SWY.CS5.ANALSYIS.SASDS (.HEMOPH)
NTT1SWY.CS5.RTI.ORI.VARS.SASDS(.HEMPH)
NTT1SWY.CS5.LIVER.DEC98.SSD(.HEMOPH)
NTT1SWY.CS5.HCVGENOS.SSD(.HCV)
NTT1SWY.HCV.COMBINED.SSD(.HCV)
NTT1SWY.CS5.SEROTYP.SSD(.SERO)
UAT1PIV.HIV9.PCR.SASDS(.PCR)

Variables

AGE0 - Age at start date
AGE3 - Age0 in terciles (values 1-3)
AGE5 - Age0 in quintiles (values 1-5)
BDNA - bdna value chosen
BIRTH - Birth date (SAS)
CENSORDT - Earlier of Liver Disease, LKA, 5/28/98
used in all analysis with Liver Disease as outcome
CNTL1 - Serotype band

CNTL2 - Serotype band
CORE1 - Serotype band
CORE2 - Serotype band
DEAD - Death flag 0/1
ENDDATE - Earlier of LKA, 5/28/98

used in all analysis with Death as outcome
F_BDNA - Subjects 1st bDNA count - for subjects with 2+ bDNA's
F_DATE - Date of subjects 1st bDNA count
GENOTYPE - HCV genotype
GROUP - Birth cohort/plus/everyone else ****
ISFROM - 'US' or 'EU'
L_BDNA - Subject's last Bdna count - for subjects with 2+ bDNA's
L_DATE - Date of subject's last bDNA count
LASTFRST - log10 of last bdna minus log10 of first bdna
LINKCD4 - CD4 count closest to the 'linked' draw date for subjects with bdna, serotyping and genotyping on same 'linked' sample
LINKCD8 - CD8 count closest to the 'linked' draw date for subjects with bdna, serotyping and genotyping on same 'linked' sample
LIVER- Liver disease 0/1
LIVERDT - Liver disease diagnosis date (SAS)
LKA - Last known alive (SAS)
NEWTYP - Hemophilia type (see AIDS file documentation)
NS41 - Serotype band
NS42 - Serotype band
NS43 - Serotype band
SAMPDRAW Draw date for samples with bdna and geno/serotypes. ie: "linked sample
SEROTYP - Selected HCV serotype
SEV - Hemophilia Severity
STARTCD4 - CD4 count closest to start date
STARTCD8 - CD8 count closest to start date
STARTDT - Start date. Later of 5/19/82 or birth (SAS)
STATANL - HIV Status
STCD4_DT - Date of 'STARTCD4' (see above)
STCD8_DT - Date of 'STARTCD8' (see above)
STUDY_ID - Study Id
SUBJ_ID - Subject Id

PDS of Mainframe programs

NTT1SWY.HCV.PAPER99.SAS

Contains all the programs written for and run on the mainframe (Kaplan Meires ect.)

'CREANAL' - copy of the program that created the analysis file

Other Files

There are 5 files in &NTT1SWY.HCV.PAPER99.SSD in addition to the analysis file "HCV"

LABDATA

&NTT1SWY.HCV.PAPER99.SSD(.LABDATA)
C:\HCV\PAPER99\LABDATA.SD2

Contains laboratory and clinical information (such as alcohol) on the subjects

Programs to create file
&NTT1SWY.HCV.PAPER99.SAS(CRELAB)
c:\hcv\paper99\CRELAB.SAS

This program did not write to the PAPER99 PDS. The day after Fran created the file 1/27/99, 1/28/99 it was copied onto

LIVER

&NTT1SWY.HCV.PAPER99.SSD(.LIVER)
C:\HCV\PAPER99\NEWLIVER.SD2

Created by Fran Yellin's program and copied into the paper99 PDS and directory at the time of analysis 1/99

RIBA

&NTTISWY.HCV.PAPER99.SSD(.RIBA)
C:\HCV\PAPER99\RIBA.SD2

Contains the RIBA data used in this analysis. Uses the RIBA results closest to the subject's STARTDT.
Program to create the file is lost

RIBAT4T8

&NTT1SWY.HCV.PAPER99.SSD(.RIBAT4T8)

C:\HCV\PAPER99\RIBAT4T8.SD2
RIBA related CD4 and CD8 data

Program NTT1SWY.HCV.PAPER99.SSD(RIBAT4T8)
C:\HCV\PAPER99\RIBAT4T8.SAS

Contains the draw date of the RIBA used in analyis, and the draw dates and CD4/CD8 counts and percentages of the CD4/CD8 closest to RIBA

Variables

RIBADRAW	-	Draw date of RIBA
RIBANT4	-	CD4 count closest to RIBA
RIBANT8	-	CD8 count closest to RIBA
RIBAPT4	-	CD4 % closest to RIBA
RIBAPT8	-	CD8 % closest to RIBA
SUBJ_ID	-	Subject Id
T4DRAW	-	Draw date of CD4
T8DRAW	-	Draw date of CD8
TOCD4	-	Time from CD4 to RIBA
TOCD8	-	Time from CD8 to RIBA

TCELLS

NTT1SWY.HCV.PAPER99.SSD(.TCELLS)
C:\HCV\PAPER99\TCELLS.SD2

The T-Cell file as it existed when analysis was begun 1/20/99 - all records for the hemophiliacs with either a CD4 or CD8 result

Dr. Eyster's Analysis for JID paper 09/98

File NTT1SWY.HCV.EYSTANAL.SSD(.HCV)
c:\hcv\hemfiles\eystanal.sd2

One record per subject/draw date, sorted by subject/draw

About the file

In September 1998 Dr. Eyster submitted a paper on the original genotype analysis she had done, on the 17/17 HIV+HIV- subjects she has been following for years. We have complete data on

16 pairs, so only 32 subjects were included in the analysis. Her analysis used both data from the 'original' HCV genotypes (date stamp 1/1/97) on the genotype file, and some additional data from Dr. Hatzakis that was new to us, but not to her.

A file was created containing the data used for this analysis. Two subject/draw dates have different genotypes on this file and the genotype analysis file. HER040 (01/08/86), and HER243 (12/17/86) were tested twice. The genotypes on this file reflect the first results that came in on them in 1997 since all her analysis is based on a batch being tested as a unit.

Input Files c:\hcv\hemfiles\allbdna.sd2
(NTT1SWY.CS5.HCVGENOS.SSD(.HCV)
C:\hcv\eyster\dataused.dat
(NTT1SWY.HCV.EYSTER.LINKUP.DAT)

Variables

ANY0	-	Genotype result is 'NEGATIVE'
ANY1	-	At least 1 Genotype 1
ANY2	-	At least 1 Genotype 2
ANY3	-	At least 1 Genotype 3
ANY4	-	At least 1 Genotype 4
ANY5	-	At least 1 Genotype 5
ANY9	-	Genotype result is POSITIVE but UNSPECIFIED
DRAW	-	Draw date (SAS)
GENOTYPE	-	Genotype
GROUP	-	Genotype group
		1 = More than 1 genotype
		2 = Negative
		3 = Positive - not otherwise specified
		4 = Any 1
		5 = Any 2
		6 = Any 3
		7 = Any 4
		8 = Any 5
GROUPTXT	-	Probably silly to include, but comes in handy in printouts
		More than 1
		Neg
		'+' NOS
		ANY 1
		ANY 2
		ANY 3
		ANY 4
		ANY 5
HCV_STAT	-	HCV Status at time of analysis
LIVFLAG	-	Liver Disease Flag

LIVERDT - Liver Failure Date (SAS)
PTNUM - Patient number Dr. Eyster uses to identify subjects
RFLP - RFLP result from Dr. Simmonds
STATANL - HIV Status

******** HER DRAW DATES ARE OFTEN OFF BY A DAY OR SO. This was attempted to be cleaned up, but the two close-in-time draw dates are often both on the background file with different sample id's. There is no pattern of which ones contains the RFLP and/or Bdna results. Therefore, any additional data on these subjects will have to be looked at carefully before analysis.

Updating the file

Run c:\hcv\hcvsas\eystanal.sas - it creates and uploads the file
Program also saved as &NTT1SWY.HCV.EYSTER.SAS(CREANAL)

HCV Serotyping File

File NTT1SWY.HCV.SEROTYP.SSD(.SERO)

One record per sample, sorted by sample.

About the file

This file contains the results from the Chiron Serotying performed at SAIC. The valid results sent from the lab consist of serotypes '1', '2',or '3', serotype 1 or 3 read as '13', and two additional codes 'UTY' and 'DIL' which stand for 'untypable' samples, and the recomondation to 'DILUTE' the sample and test again.

The program does the following. The controls are deleted. The following hierarchy is applied to the dilution results: 33, 132, 660 and 10, UTY, DIL - with 33 being the preferred dilution. One record per sample is put out. Samples that only have result/results of 'UTI' have serotype changed to 'WEAK'.

Input Files UAT1PIV.HCVSER.SSD(.SERO)

Variables
CNTL1 - Cntrl 1
CNTL2 - Cntrl 2
CORE1 - Core1 Band4
CORE2 - Core2 Band5

DILUTION	-	Dilution
DRAW	-	Draw date (SAS)
NS41	-	NS4 Band1
NS42	-	NS4 Band2
NS43	-	NS4 Band3
SAMPL_ID	-	Sample Id
SEROTYP	-	Sero Type
SUBJ_ID	-	Subject Id
WANTDIL	-	Dilution ranking (see 'About the File)

Updating the file

Run c:\hcv\hcvsas\cresero.sas - it created and uploads the file
Program also saved as &NTT1SWY.CS5.HCV.SUBTYPE.SAS(CREATE)

First Hepatitis C Sub-Genotype File

Analysis File : NTT1SWY.CS5.HCV.SUBTYPE.SASDS

Input Files: NTT1SWY.CS5.LVRRNA.MEAN.APR94
NTT1SWY.CS5.LIVERDIS.SUBTYPE.L42

Create Program: NTT1SWY.CS5.HCV.SUBTYPE.SAS(CREATE)

Dr. Hatzakis was sent samples for HCV subtypings testings. The results were put in a flat file NTT1SWY.CS5.LIVERDIS.SUBTYPE.L42. The following documatation is describing the SAS date created from this. This file contains all the variables that have been used for the preliminary analysis. The sample XA 2473 has had the draw date in the Liver Disease File (CS5.LVRRNA.MEAN.APR94) corrected to match the background file.

Issues and Recodes

A. Several formats of the HCVRNA levels are included.

1. HCVRNA_A is the adjusted RNA level (RNA*0.073)

2. RNAMULT is the adjusted RNA value multiplied by a factor based on the corresponding subtype: If the subtype is one of the '2's the HCVRNA_A is multiplied by 3. If the subtype is one of the '3's the HCVRNA_A is multipled by

B. The pos/neg PCR_EYE variable used for the original Liver Disease analysis for Dr. Eyster is included.

C. The subtyping variables based on draw date filled in and a variable that reflects the change in a subject's subtyping over time is created. The following rules apply.

1. Sub1, sub2, sub3 are filled in based on the draw date of the subtyping, even if some of the samples selected did not get tested; Therefore a person could have sub2 filled in with sub1 blank.

2. The subtype results 'missing' and 'NEG' are ignored in the creation of the 'change over time' variable.

The 'change over time' values consist of changes from and to '1A' and '3A'. The values at this time are:

1A --> No Later
1A --> Other
1A --> 1A
3A --> NO LATER
3A --> OTHER
3A --> 3A
OTHER --> DIFFER
OTHER --> NO LATER
NO RESULTS

*** The trend of OTHER --> SAME did not occur, but could with more results.

Variables:

CATS9	Char 20	Subtype Change
DRAW	Numeric	Draw Date (SAS)
HCVRNA	Numeric	Orig HCVRNA Level
HCVRNA_A	Numeric	HCVRNA * 0.073
PCR_EYS	Numeric	0/1 PCR Flag - Dr. Eyster
RNAMULT	Numeric	Adjusted * 2 or 3
SAMPL_ID	Char 7	Sample
SUB1	Char 5	Result of first draw sent
SUB2	Char 5	Result of second draw sent
SUB3	Char 5	Result of third draw sent
SUBJ_ID	Char 11	Subject
SUBTYPE	Char 5	Subtype

HGV Couples Analysis Files

About the files

706 samples from 353 spouse pairs were sent to Dr. Harvey Alter's lab for PCR and antibody testing. We have received 405 (406?) results.

The antibody flag sent from the lab is generally based on the premise that a cutoff over 0.85 is positive. There are however three samples with a cutoff over 0.85 and the antibody 0. This was done at the lab on purpose because of evidence that it was a false positive. Two new antibody flags were created for the data. High Antibody Flag is set if the cutoff is > 1.02. Low Antibody Flag is set if the cutoff is > 0.75.

&NTT1SWY.HGV.ANALYSIS.SASDS(.HCG)
SUBJECT FILE One record per sample (subject)
706 records - but not all have been tested for HGV
If tested the HGVSTAT will not be missing

&NTT1SWY.HGV.FEMHEM.SSD(.HGV)
SPOUSE PAIR FILE One record per spouse pair
353 records - but not all have been tested
If female has been tesed FEMRSLT will equal 1
If male has been tested HEMRSLT will equal 1

Input files

NTT1SWY.CS5.TOHGV.PLASMA.D0297
C:\HGV\PSENT.DAT
Plasma samples sent for testing

&NTT1SWY.CS5.TOHGV.SERUM.D0297
C:\HGV\SSENT.DAT
Serum samples sent for testing

&NTT1SWY.HGV.PLASMA.DAT.L23
C:\HGV\PLASMA.DAT
Only plasma samples with test results

&NTT1SWY.HGV.SERUM.DAT.L23
C:\HGV\SERUM.DAT
Only serum samples with test results

&NTT1SWY.CS5.ANALSYIS.SASDS(.HEMOPH)
&NTT1SWY.CS5.HCVSTAT.SASDS(.STAT)

Variables on Subject File

AGE	Age at draw
BIRTH	Birthdate (SAS)
COMBSTAT	Flag combination (see variable note **)
CUTOFF	Laboratory cutoff. >.85 consirderd positive
DRAW	Draw date (SAS)
DT_STAMP	File Date - date result put in the file
FAMILY	Family Code
HCV_STAT	HCV status
HGVAB	HGV Antibody flag
HGVSTAT	Overall HGV Status
HIGHAB	Flag indicating cutoff > 1.02
INHIB	Inhib
LABNUM	Lab #
LOWAB	Flag indicating cutoff > .75
MAT_CODE	Material B1-Serum, B2 Plasma
OD	Laboratory OD
PCR	HGV PCR flag
SAMPL_ID	Sample
SEX	Sex
STATANL	HIV Status
SUBJ_ID	Subject
TITER	log10 Titer of RNA

Variable note **

The 'combination flags' show which one (if any) of the HGV results - PCR and antibody are positive. The variable contents are 'AB', 'PCR', 'BOTH', 'NONE'. 'NONE' means that we have an antibody and PCR result and they are both negative. These flags are not set to 'NONE' to indicated missing data.

Variables on Spouse Pair File

FAMILY	Family code - matching to SUBJECT FILE
FEMAB	Female partner antibody flag
FEMAGE	Female partner age at draw
FEMBIRTH	Female partner birthdate (SAS)
FEMCOMB	Female partner flag combination (see variable note **)
FEMHCV	Female partner HCV Status

FEMHIGH	Female partner High Ab > 1.02
FEMHIV	Female partner HIV
FEMID	Female partner Id
FEMLOW	Female partner Low Ab > .75
FEMMAT	Female partner Material
FEMPCR	Female partner PCR flag
FEMRSLT	Female has result
FEMSAMP	Female partner Sample
FEMSEX	Female partner Sex
FEMTITER	Female partner HGV titer
FGSTAT	Female partner HGV
HEMAB	Hem antibody flag
HEMAGE	Hem age at draw
HEMBIRTH	Hem birth date (SAS)
HEMCOMB	Hem flag combination (see variable note **)
HEMHCV	Hem HCV status
HEMHIGH	Hem High antibody flag 1.02
HEMHIV	Hem HIV Status
HEMID	Hem Id
HEMLOW	Hem Low Ab > .75
HEMMAT	Hem Material
HEMPCR	Hem PCR
HEMRSLT	Has Result
HEMSAMP	Hem Sample
HEMSEX	Hem Sex
HEMTITER	Hem HGV titer
HIVPAIR	HIV Pair Status
MFPAIR	HGV Pair Status
MGSTAT	Hem HGV Status

Updating the Files

C:\HGV\CREANAL.SAS	&NTT1SWY.CS5.HGV.SAS(CREANAL)